Chapter 17
Functional Two-sample Tests Based on Empirical Characteristic Functionals

Zdeněk Hlávka and Daniel Hlubinka

Abstract Two-sample tests for functional data based on empirical characteristic functionals are proposed. The test statistic is of Cramér–von Mises type with integration over a preselected family of probability measures, say Q, leading a computationaly feasible and powerful test statistic. The choice of the probability measure Q is discussed and the empirical size and power of the resulting two-sample functional tests are investigated in a small simulation study.

17.1 Introduction

Functional data analysis already became a standard [11, 6, 7, 9] with many tools obtained as a generalization of a corresponding multivariate method. In this contribution, we investigate the general functional two-sample problem and propose a new two-sample functional test statistic based on empirical characteristic functionals.

Assuming two functional random samples, say $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$, the problem of testing the null hypothesis of equality of the respective mean functions, i.e.,

$$H_0 : m_X(.) = m_Y(.)$$

has already been extensively investigated, see [3] for an overview. Slightly different hypothesis is studied in [8], namely

$$H_0 : \forall_t \; X(t) =_{\mathcal{L}} Y(t)$$

Zdeněk Hlávka
Univerzita Karlova, Dept. of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Sokolovská 83, Praha 8, Czechia, e-mail: hlavka@karlin.mff.cuni.cz

Daniel Hlubinka (✉)
Univerzita Karlova, Dept. of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Sokolovská 83, Praha 8, Czechia, e-mail: hlubinka@karlin.mff.cuni.cz

G. Aneiros et al. (eds.), *Functional and High-Dimensional Statistics and Related Fields*, Contributions to Statistics, https://doi.org/10.1007/978-3-030-47756-1_17

testing simultaneously the distribution of all projections where $=_{\mathcal{L}}$ denotes the equality in distribution.

In the following, instead of comparing only the mean functions or testing the distributions of projections, we are interested in testing a more general null hypothesis of equality of entire functional distributions:

$$\mathcal{H}_0 : \phi_X = \phi_Y \tag{17.1}$$

where ϕ_X and ϕ_Y denotes, respectively, the *characteristic functional* (CF) of the X and Y sample. The definition and properties of CF and *empirical CF* (ECF) are summarized in Section 17.2. A two-sample test statistic based on a distance between two ECFs is proposed in Section 17.3. Finally, a small simulation study in Section 17.4 investigates small sample properties of the ECF-based two-sample test.

17.2 Empirical Characteristic Functional

In what follows, we consider functional random variables with values in the space of continuous functions or in the space of measurable square integrable functions, i.e., $X : \Omega \to C[0,1]$ or $X : \Omega \to \mathcal{L}_2[0,1]$, where the domain is as usually (and wlog) chosen to be $[0,1]$.

The CF of X is $\phi_X(u) = \mathrm{E}\exp(i\langle u, X\rangle)$ for $u \in C^*[0,1]$ or $u \in \mathcal{L}_2^*[0,1]$, the dual space of $C[0,1]$ or $\mathcal{L}_2[0,1]$, respectively. Due to the properties of CF, it is sufficient to consider just $u \in \mathcal{L}_2^*[0,1] = \mathcal{L}_2[0,1]$ for both options in which case $\langle u, X\rangle = \int_0^1 u(t)X(t)\mathrm{d}t$.

The ECF of a functional random sample $X_1, \ldots, X_n$ is

$$\widetilde{\phi}_X(u) = \frac{1}{n}\sum_{k=1}^{n} \exp\big(i\langle u, X_k\rangle\big).$$

The functional data are not observed continuously in most cases. We may consider all X_i's to be observed on a regular grid of points $t_j = j/N$, $j = 0, 1, \ldots, N$ since the generalisation to different observation points is straightforward. The ECF is then

$$\widehat{\phi}_X(u) = \frac{1}{n}\sum_{k=1}^{n} \exp\big(i\langle u, X_k\rangle_d\big),$$

where $\langle u, X\rangle_d = \sum_{i=1}^{N} u(t_i)X(t_i)(t_i - t_{i-1}) = \frac{1}{N}\sum_{i=1}^{N} u(t_i)X(t_i)$.

17.3 Cramér–von Mises Type of Statistics

Our Cramér–von Mises two-sample test is based on two random samples $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ with the test statistic

$$\int_{\mathcal{L}_2[0,1]} |\widehat{\phi}_X(u) - \widehat{\phi}_Y(u)|^2 \mathrm{d}Q(u), \tag{17.2}$$

where Q is some probability measure on the space $\mathcal{L}_2[0,1]$ discussed later. The squared distance $|\widehat{\phi}_X(u) - \widehat{\phi}_Y(u)|^2$ of the ECFs may be rewritten as

$$\left(\frac{1}{n}\sum_{k=1}^{n}\cos\langle u, X_k\rangle_d - \frac{1}{m}\sum_{\ell=1}^{m}\cos\langle u, Y_\ell\rangle_d\right)^2 + \left(\frac{1}{n}\sum_{k=1}^{n}\sin\langle u, X_k\rangle_d - \frac{1}{m}\sum_{\ell=1}^{m}\sin\langle u, Y_\ell\rangle_d\right)^2 \tag{17.3}$$

and we obtain after some calculation and using the trigonometric identity the final form

$$\frac{1}{n^2}\sum_{k,j=1}^{n}\cos\langle u, X_k - X_j\rangle_d + \frac{1}{m^2}\sum_{\ell,j=1}^{m}\cos\langle u, Y_\ell - Y_j\rangle_d - \frac{2}{mn}\sum_{k=1}^{n}\sum_{\ell=1}^{m}\cos\langle u, X_k - Y_\ell\rangle_d. \tag{17.4}$$

17.3.1 Choice of Q

The measure Q is some probability measure on the space $\mathcal{L}_2[0,1]$. We propose to use a special form of this measure, namely some special form of a Gaussian measure. Hence, we consider a random function U with all finite-dimensional distribution being multivariate normal distribution. Since the data are observed on a discrete grid of N points, it is sufficient to consider a random vector $U_N = \big(U(t_1), \ldots, U(t_N)\big)$ following zero mean N-dimensional normal distribution with the variance matrix $\boldsymbol{V} = (v_{i,j})_{i,j=1}^N$. For a fixed discretely observed function x, we have

$$\langle U, x\rangle_d = \frac{1}{N}\sum_{i=1}^{N} U(t_i)x(t_i) \sim \mathcal{N}\left(0, \frac{1}{N^2}\sum_{j,k=1}^{N} x(t_i)x(t_j)v_{i,j}\right) = \mathcal{N}\big(0, \sigma^2(x)\big),$$

where $\sigma^2(x) = \frac{1}{N^2}x^T \boldsymbol{V} x$ and $\mathrm{E}_Q \cos\langle U, x\rangle_d$ becomes

$$\mathrm{E}_Q \cos\left(\frac{1}{N}\sum_{i=1}^{N} U(t_i)x(t_i)\right) = \exp\left(-\frac{1}{2}\sigma^2(x)\right) \tag{17.5}$$

and the test statistic (17.2) based on the two samples $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ becomes

$$
\begin{aligned}
T = & \frac{1}{n^2} \sum_{k,j=1}^{n} \exp\left(-\frac{1}{2N^2}(X_k - X_j)^T \boldsymbol{V}(X_k - X_j)\right) \\
& + \frac{1}{m^2} \sum_{k,j=1}^{m} \exp\left(-\frac{1}{2N^2}(Y_k - Y_j)^T \boldsymbol{V}(Y_k - Y_j)\right) \\
& - \frac{2}{nm} \sum_{k=1}^{n} \sum_{j=1}^{m} \exp\left(-\frac{1}{2N^2}(X_k - Y_j)^T \boldsymbol{V}(X_k - Y_j)\right).
\end{aligned}
\tag{17.6}
$$

The null hypothesis will be rejected for large values of the test statistic T, i.e., for

$$
T \geq c(\alpha), \tag{17.7}
$$

where $c(\alpha)$ denotes critical value such that $P(T \geq c(\alpha)|\mathcal{H}_0) = \alpha$. In the following, the critical value will be approximated by the permutation principle [2].

17.3.2 The Matrix *V*

The performance of the test largely depends on the matrix $\boldsymbol{V}$ introduced in Section 17.3.1. We propose several possibilities and our test is then compared with other two-sample tests in a small simulation study.

The most simple choice is to set $\boldsymbol{V} = \mathcal{I}_N$ but the following possibilities should have better power.

Variance matrix of a Gaussian process: This proposal follows classical "random projection" approach. It is considered that U is a Gaussian process, usually a Wiener process and $\boldsymbol{V} = \Sigma_W$ is the variance matrix of the process observed at $j/N, j = 1, 2, \ldots, N$.

The observations: We consider $n + m$ iid $Z_\ell \sim \mathcal{N}(0, 1)$, and

$$
U = \frac{1}{\sqrt{n+m}} \left[\sum_{j=1}^{n} Z_j X_j + \sum_{k=1}^{m} Z_{k+n} Y_k \right].
$$

Then

$$
z_{q,r} = \frac{1}{N^2(n+m)} \left[\sum_{j=1}^{n} X_j(t_q) X_j(t_r) + \sum_{k=1}^{m} Y_k(t_q) Y_k(t_r) \right]
$$

and we can set $\boldsymbol{V} = \boldsymbol{Z} = (z_{q,r})_{q,r=1,\ldots,N}$.

Sample covariance matrix: By centering the (functional) observations, we actually obtain $\boldsymbol{V} = \hat{\Sigma}$, where $\hat{\Sigma}$ denotes the sample variance matrix of the observed N-dimensional random vectors (approximating the functional observations).

Notice that the quadratic forms in exponential functions in (17.6) look similarly to Hotelling's T^2 test statistic, where the matrix $\boldsymbol{V}$ is chosen as the inverse of the sample covariance matrix. Therefore, further possible choices of the matrix $\boldsymbol{V}$ could be the inverse of the matrix $\hat{\Sigma}$ or Σ_W. Note that the inverse of $\hat{\Sigma}$ generally does not exist

but, depending on the number of observations, first d eigenvectors and eigenvalues can be used to calculate a simple approximation. Interestingly, the eigenvectors (or eigenfunctions) tend to recover the direction (in the functional space) that separates the two sets of functional observations, see also the discussion in [5].
Eigenvectors and eigenvalues: Denote by $\lambda_1 \geq \lambda_2 \geq \dots$ the ordered eigenvalues, and by $e_1, e_2, \dots$ the corresponding orthogonal eigenfunctions of the covariance operator of the combined dataset. Consider $d \geq 1$ and iid random variables $Z_\ell \sim \mathcal{N}(0,1)$, and define

$$U = \sum_{\ell=1}^{d} \frac{1}{\sqrt{\lambda_\ell}} e_\ell Z_\ell.$$

The theoretical eigenfunctions and eigenvalues are replaced by eigenvectors $\hat{e}_\ell$ and eigenvalues $\hat{\lambda}_\ell$ of the empirical variance matrix $\hat{\Sigma}$ in practical applications. Then for some $1 \leq d \leq \min(m+n, N)$ define

$$e_{q,r} = \sum_{\ell=1}^{d} \frac{1}{\hat{\lambda}_\ell} \hat{e}_\ell(t_q)\hat{e}_\ell(t_r).$$

In the following, we denote the resulting matrix $\boldsymbol{V} = (e)_{q,r=1,\dots,N} = \hat{\Sigma}_d^{-1}$. The choice of d is discussed later.

17.4 Simulation and Comparison

We start by investigating the empirical power against the 'location shift' alternative. Following [4, Section 5], we generate two functional samples

$$X_i(t) = \mu_x(t) + \varepsilon_{x,i}(t) \tag{17.8}$$

and

$$Y_i(t) = \mu_y(t) + \varepsilon_{y,i}(t),$$

where the mean functions are $\mu_x(t) = (1, 2.3, 3.4, 1.5)(1, t, t^2, t^3)^\top$ and $\mu_y(t) = \mu_x(t) + 2\delta(1, 2, 3, 4)(1, t, t^2, t^3)^\top/\sqrt{30}$ so that the parameter δ controls the difference between $\mu_x(t)$ and $\mu_y(t)$. The subject-effect functions $\varepsilon_{\cdot,i}(t)$ are defined as a random linear combination of 11 orthonormal basis vectors $\psi_w(t)$ (such that $\psi_1(t) = 1$, $\psi_{2\omega}(t) = \sqrt{2}\sin(2\pi\omega t)$, $\psi_{2\omega+1}(t) = \sqrt{2}\cos(2\pi\omega t)$, for $\omega = 1, \dots, 5$) with coefficients $b_{\cdot,i,w} \sim N(0, 1.5\rho^w)$, for $w = 1, \dots, 11$.

In this section, we set $\rho = 0.5$. The choice $\delta = 0$ means that the null hypothesis is satisfied and we investigate the empirical size. An example of two data sets generated under the alternative, with $\delta = 0.5$, is plotted in Figure 17.1. Note that these two samples were generated in the same way as samples 2 and 3 in [4, Section 5.2].

We compare empirical sizes and powers of the proposed two-sample ECF-based test (ECF) with various variance matrices $\boldsymbol{V}$, described in Section 17.3.2, to tests implemented in R library fdANOVA [4]. Many of these tests are based on the usual

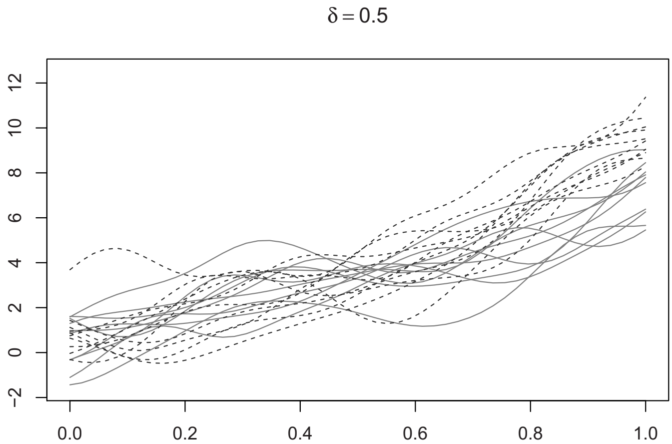

Fig. 17.1 Random samples $X_1, \ldots, X_{10}$ (solid lines) and $Y_1, \ldots, Y_{10}$ (dashed lines) generated according to the algorithm described in Section 17.4 with $\delta = 0.5$ and $\rho = 0.5$.

univariate (pointwise) F-statistics, say $F_n(t)$ for $t \in (0, 1)$, that are combined into a single test statistic:

GPF: globalizing pointwise F test, $T^{\text{GPF}} = \int F_n(t)dt$,
Fmaxb: maximizing pointwise F test, $T^{\text{Fmaxb}} = \max F_n(t)$,

Another approach is based on testing k projections of the original functions by combining p-values [1] that are based on:

ANOVA: ANOVA F-test statistic,
ATS: ANOVA-type statistic,
WTPS: Wald type permutation statistic.

Similarly to the choice of the random process (and matrix $\boldsymbol{V}$) in Section 17.3.2, the projections are generated either as Gaussian white noise (G) or Brownian motion (B). A detailed description of the function `fanova.tests()` in R library `fdANOVA` is given in [4].

In the first two columns of Table 17.1, we can see that the empirical size of all tests is close to the nominal level $\alpha = 5\%$ both for $n = m = 10$ and $n = m = 20$ observations.

For $n = m = 10$, the empirical power is smallest for ECF tests with $\boldsymbol{V} = \Sigma_W^{-1}$ (7.8%) and $\boldsymbol{V} = \boldsymbol{Z}$ (27.6%). The empirical power of most tests lies between 47% and 60%. Somewhat higher power, almost 70%, has been obtained for Fmaxb and ECF tests with $\boldsymbol{V} = \hat{\Sigma}_6^{-1}$. Using $d = 8$ eigenvectors, the highest empirical power (90.5%) is observed for the ECF test with $\boldsymbol{V} = \hat{\Sigma}_8^{-1}$.

		$\delta = 0$		$\delta = 0.5$	
method	test/$\boldsymbol{V}$	$n = 10$	$n = 20$	$n = 10$	$n = 20$
GPF		6.1	6.2	58.0	88.4
FMAXB		6.1	3.6	69.8	98.0
P-Gauss	ANOVA	3.5	3.3	51.1	90.1
	ATS	5.0	4.9	54.2	89.6
	WTPS	3.8	3.6	47.8	90.5
P-BM	ANOVA	2.7	3.4	54.7	87.7
	ATS	3.4	3.5	55.5	87.0
	WTPS	2.6	3.1	47.5	88.0
ECF	$\mathcal{I}$	4.7	4.0	53.9	89.3
	Σ_W	4.4	4.8	57.5	87.8
	Σ_W^{-1}	5.8	4.8	7.8	11.8
	$\boldsymbol{Z}$	4.8	4.9	27.6	52.1
	$\hat{\Sigma}$	4.5	5.6	43.2	76.6
	$\hat{\Sigma}_2^{-1}$	4.2	5.7	47.6	79.4
	$\hat{\Sigma}_6^{-1}$	5.2	5.9	67.6	95.2
	$\hat{\Sigma}_8^{-1}$	4.2	4.9	**90.5**	**99.9**

Table 17.1 Empirical size ($\delta = 0$) and empirical power ($\delta = 0.5$) (in %) of two-sample functional tests, nominal level $\alpha = 0.05$, $\rho = 0.5$, $N = 50$ gridpoints, equally sized samples ($n = m$), 1000 simulations with 1000 permutations. Bold font denotes the highest observed empirical power.

Results for $n = m = 20$ are similar but observed differences are smaller because the power of most tests is close to 90%.

	F-statistic			ECF		
		$n = 10$	$n = 20$	$\boldsymbol{V}$	$n = 10$	$n = 20$
GPF		7.6	5.4	$\mathcal{I}$	6.3	9.2
FMAXB		6.9	6.2	Σ_W	22.4	54.1
P-Gauss	ANOVA	4.7	4.4	Σ_W^{-1}	8.5	14.8
	ATS	4.9	3.3	$\boldsymbol{Z}$	28.6	59.1
	WTPS	4.9	4.5	$\hat{\Sigma}$	**53.2**	**89.5**
P-BM	ANOVA	3.5	3.1	$\hat{\Sigma}_2^{-1}$	6.4	4.9
	ATS	3.6	3.9	$\hat{\Sigma}_6^{-1}$	9.8	7.0
	WTPS	3.1	3.0	$\hat{\Sigma}_8^{-1}$	10.0	10.6

Table 17.2 Empirical power ($\sigma_y = 2$) (in %) of two-sample functional tests, nominal level $\alpha = 0.05$, $\rho = 0.5$, $N = 50$ gridpoints, equally sized samples ($n = m$), 1000 simulations with 1000 permutations. Bold font denotes the highest observed empirical power.

The study of the empirical power against the 'change-of-scale' alternative is summarized in Table 17.2; the random functions $X_i(t)$ are still generated according to (17.8) while the second sample is changed to

$$Y_i^{\sigma}(t) = \mu_y(t) + \sigma_y \varepsilon_{y,i}(t),$$

with $\delta = 0$ (implying that $\mu_x(.) = \mu_y(.)$) and with additional parameter $\sigma_y > 0$ controlling the variance. As may be expected, the empirical power of the F-statistic-based tests shown in Table 17.2 is very close to the nominal test level. The best power is obtained for the ECF-based test with $\boldsymbol{V} = \hat{\Sigma}$, the sample covariance matrix. On the other hand, the ECF-based tests using $\boldsymbol{V}$ based on the inversion of (some) covariance matrix do not perform very well.

We conclude that the ECF test with the matrix $\boldsymbol{V}$ approximating the inverse covariance matrix leads to the best results against the 'location shift' alternative while the ECF test with $\boldsymbol{V} = \hat{\Sigma}$ leads to the best results against the 'change-of-scale' alternative. Interestingly, the ECF test outperforms the F-statistic-based tests implemented in library `fdANOVA` even against the 'location shift' alternative.

Acknowledgements The research was supported by Czech Science Foundation, grant GAČR 18-08888S.

References

[1] Cuesta-Albertos, J.A., Febrero-Bande, M.: A simple multiway ANOVA for functional data. Test **19**, 537–557 (2010)
[2] Good, P.: Permutation, Parametric, and Bootstrap Tests of Hypotheses. Springer-Verlag, New York (2005)
[3] Górecki, T., Smaga, Ł.: A comparison of tests for the one-way ANOVA problem for functional data. Comp. Stat. **30**, 987–1010 (2015)
[4] Górecki, T., Smaga, Ł.: fdANOVA: an R software package for analysis of variance for univariate and multivariate functional data. Comp. Stat. **34**, 571–597 (2019)
[5] Hall, P., Poskitt, D.S., Presnell, B.: A functional data-analytic approach to signal discrimination, Technometrics **43**, 1–9 (2001)
[6] Horváth, L., Kokoszka, P.: Inference for Functional Data with Applications. Springer (2012)
[7] Hsing, T., Eubank, R.: Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators. Wiley (2015)
[8] Jiang, Q., Hušková, M., Meintanis, S.G., Zhu, L.: Asymptotics, finite-sample comparisons and applications for two-sample tests with functional data. Journal of Multivariate Analysis **170**, 202–220 (2019)
[9] Kokoszka, P., Reimherr, M.: Introduction to Functional Data Analysis. Chapman and Hall/CRC (2017)
[10] Meintanis, S.G.: A review of testing procedures based on the empirical characteristic function. South African Statistical Journal **50**, 1–14 (2016)
[11] Ramsay, J.O., Silverman, B.W.: Applied Functional Data Analysis: Methods and Case Studies. Springer (2007)